Assignment

Due on Wednesday, September 11, 2019

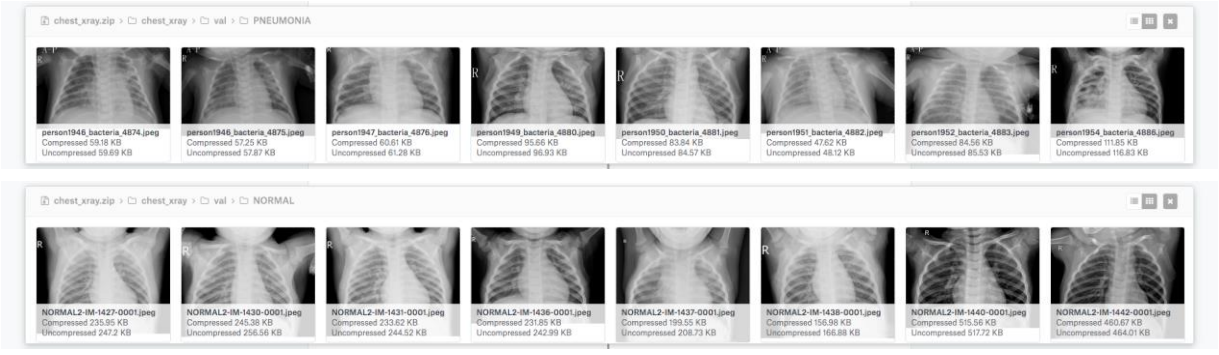
Advanced Lectures on Pattern Processing
by Tomoya Sakai

You will write a report on chest X-ray classification experiments in the following senario, using sample codes with tensorflow.

EXPERIMENTS

- Aims
  - Show your understanding on how to build a convolutional neural network (CNN) classifier.
  - Explain why the resulting classifier is reasonably tuned by providing convincing reasons/evidence.
- Senario
  - The task is chest radiography: input and output of the classifier are respectively a chest X-ray image and probabilities of being normal and pneumonia. The experiments use chest X-ray images, which is open to the public as

    Kermany, *et al.*, "Labeled optical coherence tomography (OCT) and chest X-ray images for classification", Mendeley Data, Version 2, 2018. (ChestXRay2017.zip)
    DOI: http://dx.doi.org/10.17632/rscbjbr9sj.2#file-41d542e7-7f91-47f6-9ff2-dd8e5a5a7861



https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

  - To build and evaluate the CNN classifier, you will undertake two practices, namely, Part I (mandatory) and Part II (optional), as described below. Sample codes in Jupyter Notebook (ipynb) are available from

    https://github.com/tsakailab/prml/tree/master/ChestXray

    On the LACS course website, you can find the links to open the sample codes in Google Colabratory.

- Part I: Design Your Own CNN (mandatory)
  - ✧ **Dataset:** The sample code `ChestXray_CCN_bbXXXXXXXX.ipynb` uses a resized version of the chest X-ray images.
    - ☐ Write a summary of the dataset in your report.
  - ✧ **CNN modeling and training:** Complete `ChestXray_CCN_bbXXXXXXXX.ipynb` to build your own CNN classifier. Read the code and comments to see how it works and what to do.
    - ☐ Describe the CNN architecture and explain why you designed it as such.
    - ☐ When fitting the CNN model to the training dataset, how was the convergence of the training loss and accuracy? How did you terminate the training?
  - ✧ **Evaluation:** Observe the predictions and visualized feature maps.
    - ☐ Report the prediction performance (not only the test accuracy but precision and recall).
    - ☐ What did (should) you do to achieve a good performance? Show why as well.
    - ☐ What do the feature maps tell you?
  - ✧ **Private scoring:** You can save the CNN model in h5 format at the end of the code. Then, open `ChestXray_PrivateScore.ipynb` and upload your best CNN model.
    - ☐ Include in the report the displayed test loss and accuracy.
    - ☐ What do you think is the purpose of this private scoring?
  - ✧ **Submission of your code:** Upload your ipynb at LACS. The filename should be `ChestXray_CNN_bbXXXXXXXX.ipynb` where `XXXXXXXX` is your student ID. Do not clear the outputs of all cells. Indicate your modification clearly as comments or texts. The shorter the running time and the simpler your code, the better.

- Part II: Transfer Learning (optional)
  - ✧ Play around with `ChestXray_FCN_TL.ipynb` if you like.
    - ☐ How is transfer learning implemented?
    - ☐ What is the FCN with GAP? Explain its advantages.
    - ☐ Which pretrained model is used in the original paper, and how much accuracy was achieved by transfer learning?

REPORTING REQUIREMENTS

☐ The PDF filename of your report should be `ChestXray_bbXXXXXXXX.pdf` where `XXXXXXXX` is your student ID.

☐ Clearly written in Japanese or English.

☐ The contents fill at least four pages in A4 format including figures.

☐ Describe what each figure shows, and explain what it means.

☐ Do not contain source codes. Include the lines only if necessary.

☐ Provide convincing reasons for your definitions/settings.

☐ Try to explain why for every cause and result.

☐ Cite reliable sources and show the list of references. Note that a URL alone is not a citation.

☐ Justify if your report does not satisfy any one of these requirements.

Submission: Assignment, Advanced Lectures on Pattern Processing @LACS

Due date: Wednesday, September 11, 2019

Questions: tsakai@cis.nagasaki-u.ac.jp